Research Article

A New Empirical Model for Short-Term Forecasting of the Broadband Penetration: A Short Research in Greece

Salpasaranis Konstantinos and Stylianakis Vasilios
Electrical and Computer Engineering Department, University of Patras, Rio, 26500 Patras, Greece

Correspondence should be addressed to Salpasaranis Konstantinos, salpk@upatras.gr

Received 25 January 2011; Accepted 26 August 2011

Academic Editor: Luis Carlos Rabelo

Copyright © 2011 S. Konstantinos and S. Vasilios. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The objective of this paper is to present a short research about the overall broadband penetration in Greece. In this research, a new empirical deterministic model is proposed for the short-term forecast of the cumulative broadband adoption. The fitting performance of the model is compared with some widely used diffusion models for the cumulative adoption of new telecommunication products, namely, Logistic, Gompertz, Flexible Logistic (FLOG), Box-Cox, Richards, and Bass models. The fitting process is done with broadband penetration official data for Greece. In conclusion, comparing these models with the empirical model, it could be argued that the latter yields well enough statistics indicators for fitting and forecasting performance. It also stresses the need for further research and performance analysis of the model in other more mature broadband markets.

1. Introduction

The diffusion of innovative products is the process in which innovation is adopted by society. An innovative product, after it has been widely adopted by society, can be the basis for developing other innovations. The modern diffusion theory concerning the construction of diffusion models for product adoption has been a topic of academic interest since the 1950s. The literature on this area is significantly large. In this brief historical overview, some important papers are presented.

Most of the diffusion models, used in market studies, are S-shaped curves. S-shaped curves (sigmoids) describe the growth of various phenomena in physics, biology, and social sciences [1]. The Logistic model has been used by Griliches (1957) in explaining the adoption of hybrid corn in the USA [2]. The linear form of the model was also used by Mansfield (1961) [3, 4]. A model derived from a study of the plants growth is that of Richards (1959) [5, 6]. The theory of diffusion of new products adopted by a social system has been presented by Rogers (1962) [7].

Rogers considers five categories according to the time response of consumers in purchasing new products. These categories of consumers, profiles are: innovators, early adopters, early majority, late majority, and laggards. This study is a standard reference for diffusion theory. The theory of Rogers is a description of the life cycle of a new product [7].

The Gompertz forecasting model was proposed by Gregg, Hassel, and Richardson (1964) [8]. According to Chow (1967), the Gompertz model is better than the Logistic model in explaining computer demand [9, 10]. Bass (1969) proposes a growth model, which accurately forecasts peak color TV sales [11]. The Bass model proposes the coefficient of innovation in diffusion’s equation, which makes it suitable for the modeling of product adoption [11, 12]. Bewley and Fiebig (1988) use the flexible logistic growth (FLOG), as well as the Box-Cox model with applications in telecommunications [13, 14].

In conclusion, it can be said that the objective of diffusion models is to describe and to predict market trends [1].

2. Diffusion Models

In general, diffusion models are deterministic time functions, and they use historical data to estimate the parameters of the diffusion process of a product’s life cycle. These models are
used to forecast the cumulative adoption of a new technology [15].

Diffusion models are divided into, at least, two categories according to whether the level of saturation is constant (nondynamic models) or changing over time (dynamic models).

The differential equation which describes the fundamental diffusion model follows the following formulation:

\[ \frac{dY(t)}{dt} = [A - Y(t)] \cdot f(t), \]  \(1\)

where \(A\) is the estimated diffusion saturation level for time \(t\), \(Y(t)\) is the diffusion penetration and function \(f(t)\) is the diffusion coefficient. In dynamic diffusion models, the \(A\) saturation factor is dependent on time. It should be considered that the difference \(A - Y(t)\) corresponds to the cumulative number of potential adopters, \(N\), for time \(t\). So, \((1)\) could be written as

\[ \frac{dY(t)}{dt} = N(t) \cdot f(t). \]  \(2\)

In this paper the formulations of the Gompertz, Logistic, Richards, Flexible Logistic (FLOG), Box-Cox are used, as well as that of the Bass model. Here follows a short review of the above models.

2.1. Gompertz Model. The Gompertz model is described by the following equations:

\[ Y = A \cdot e^{B \cdot e^{-C \cdot t}}, \]  \(3\)

\[ Y = A \cdot e^{B \cdot e^{-C \cdot t}} + D, \]  \(4\)

where \(Y = Y(t)\) is the estimated diffusion level at time \(t\), \(B < 0\) and parameter \(A\) is the saturation level [16]. The parameters that determine the model are \(A, B, C\). Parameter \(B\) is related to the year that diffusion reaches the point of inflection, and parameter \(C\) measures the “diffusion speed”. Parameter \(D\) is a constant \((4)\) and is called Gompertz with constant.

2.2. Logistic Model. The general form of the logistic model is given by

\[ Y = \frac{A}{1 + B \cdot e^{C \cdot t}}, \]  \(5\)

where \(f(t) = -C - Dt(m,k)\) and \(C, D, m, k \neq 0\) constants. Various functions derive from the logistic model such as the following.

2.2.1. Fisher-Pry Model. When the parameters of the logistic model are \(C = 0\) and \(D \neq 0\) and time \((m,k) = t\), the model is known as linear logistic or Fisher-Pry model. The linear logistic model has an inflection point that occurs when the diffusion level is the half of its saturation level.

2.2.2. Flexible Logistic Model. The form of Flexible Logistic model (FLOG) derives from the logistic, where \(C \neq 0, D \neq 0\) and

\[ t(m,k) = \left[ (1 + k \cdot t)^{1/k} \right]^m - 1, \]  \(6\)

The FLOG model locates the point of inflection anywhere between its upper and lower bounds [4]. The parameters are \(m \neq 0, k \neq 0\) [16].

2.2.3. Box-Cox Model. The Box-Cox model comes from the logistic, where \(t(m,k) = ((1 + t)^m - 1)/m\). The general logistic parameters are \(m \neq 0\) and \(k = 0\) [16].

2.3. Richards Model. The Richards’ growth model introduces an additional parameter \(D\) to the linear logistic model [17]:

\[ Y = \frac{A}{(1 + B \cdot e^{C \cdot t})^D}, \]  \(7\)

where constants \(B, C, D \neq 0\).

2.4. Bass Model. The Bass model includes two categories of adopters, the innovators, at the early stage of the diffusion, and the imitators, afterwards. \(Y = Y(t)\), the diffusion level at time \(t\), is expressed as

\[ Y = A \cdot \frac{1 - e^{-(p+q) \cdot t}}{1 + ((q/p)e^{-(p+q) \cdot t})}, \]  \(8\)

where \(A\) is the saturation level of adoption, \(p \neq 0\) is the innovation coefficient, and \(q \neq 0\) is the imitation coefficient [6, 11].

3. The Empirical Diffusion Model

3.1. General Aspect of the Model. An empirical model for short-term forecasting of the broadband penetration is proposed, described by the following discrete equation of differences:

\[ Y_{n+1} = Y_n + C_n \cdot (t_{n+1} - t_n), \]  \(9\)

where \(Y_n\) is the estimated penetration level at time \(t_n\). For \(t_0 = 0\) the initial penetration \(Y_0\) is taken. \(C_n\) function measures the “adoption rate” at time \(t_n\), given by

\[ C_n = \frac{A \cdot \frac{t_n^k}{e^{B \cdot t_n} - 1}}{1 - e^{B \cdot t_n}}, \]  \(10\)
where $k \neq 0$, $k \in R$ and model’s time $t_n$ is given by:

$$t_n = D \cdot [(n - 1) + n_{STEP}],$$

(11)

where $n \in N^*$ is the current time (natural number without 0) of the observation, $n_{STEP} \in R$ is the step between $n - 1$, and $n$ and $D \in R^4$ (real numbers without 0).

Also, an equivalent mathematical expression in (9) that can be used is

$$Y_{n+1} = Y_0 + \sum_{m=0}^{n} C_m \cdot (t_{m+1} - t_m).$$

(12)

$Li_s(z)$ represents the polylogarithm function of order $s$ and argument $z$, defined by the infinite series

$$Li_s(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^s}. $$

(15)

In this paper, the discrete format of this model is used, described in (9) or (12), (10), and (11).

3.2. Model’s Analysis. The empirical model uses the formalism of Planck’s model for black body emission to introduce the parameters in macroscopic level. The parameters of the adoption rate model into (10) are $A$, $k$, $B$, and $D$. It belongs to the dynamic diffusion type of models, meaning that the cumulative number of potential adopters $N(t_n)$ is changing against time. It follows that the estimation of the saturation level is time dependant. Parameter $A$ is a scaling factor of the model that corresponds to the initial broadband market size (potential adopters).

It should be noted that the penetration growth rate is proportional to the number of potential adopters and to the likelihood $f(t_n)$ of the fact that total number of adopters is $N(t_n)$ for time $t_n$. This can be written in discrete format as

$$\frac{\Delta Y(t_n)}{\Delta t_n} = C_N = N(t_n) \cdot f(t_n),$$

(16)

where $\Delta t_n = t_{n+1} - t_n$.

Exponent $k$ relates to the behavior of the time function which describes the total number of potential adopters. According to the empirical model, $N(t_n)$ is given by

$$N(t_n) = A \cdot t_n^k,$$

(17)

The $B$ factor ties up with the speed of the cumulative adoption and the saturation level of the diffusion model. Specifically, the $B$ factor is related to the probability distribution of the model. The discrete probability $p_n$ of total potential adopters $[N(t_n)]$ at time $t_n$ is presented in

$$p_n = e^{-B \cdot t_n}.$$  

(18)

The continuous format of (9) is given by

$$\frac{dY(t)}{dt} = A \cdot t^k \cdot e^{-B \cdot t},$$

(13)

where, for $t_0 = 0$, we take the initial diffusion level $Y(t_0) = Y_0$.

Regression analysis shows that, in many cases, the range of $k$ factor varies between values 2, 3. In the special case that $k = 3$, the solution of (13) gives:

$$Y(t) = A \cdot \left[ 6 \cdot Li_4(e^{B \cdot t}) - B \cdot t \cdot Li_3(e^{B \cdot t}) + \frac{B^2 \cdot t^2 \cdot Li_2(e^{B \cdot t})}{2} + \frac{B^3 \cdot t^3 \cdot ln(1-e^{B \cdot t})}{6} - \frac{t^4}{4} \right],$$

(14)

Equation (18) is the discrete resultant of the reasonable assumption that the rate of probability in time will be in negative proportion to the probability $p_n$ and $B$ factor [11]. The continuous format of the diffusion model is given in

$$dY(t)/dt = -B \cdot p(t).$$

(19)

The normalized discrete probability function for time $t_n$ is given in

$$f(t_n) = \frac{e^{-B \cdot t_n}}{1 - e^{-B \cdot t_n}}.$$  

(20)

The factor $1/(1 - e^{-B \cdot t_n})$ is the sum of the geometric series $1/(1 - e^{-B \cdot t_n}) = \sum_{m=0}^{\infty} e^{-B \cdot t_n}$. So, (20) is equivalent to

$$f(t_n) = (e^{-B \cdot t_n}) \cdot \left( \sum_{m=0}^{\infty} e^{-B \cdot t_n} \right) = \sum_{m=1}^{\infty} e^{-B \cdot t_n},$$

(21)

Replacing (21) in (16) yields

$$C_n = N(t_n) \cdot f(t_n) = (A \cdot t_n^k) \cdot \left( \sum_{m=1}^{\infty} e^{-B \cdot t_n} \right),$$

(22)

which is equivalent to (10).

Finally, by replacing (22) in (12), the following is obtained:

$$Y_{n+1} = Y_0 + \sum_{m=0}^{n} \left( A \cdot t_m^k \right) \cdot \left( \sum_{m=1}^{\infty} e^{-B \cdot t_n} \right) \cdot (t_{m+1} - t_m)$$

$$= Y_0 + \sum_{m=0}^{n} \sum_{\mu=1}^{\infty} \left( A \cdot t_m^k \right) \cdot e^{-B \cdot t_n} \cdot (t_{m+1} - t_m).$$

(23)

The difference $t_{m+1} - t_m$ is a constant number and equal to $D$. So, (23) becomes

$$Y_{n+1} = Y_0 + D \cdot \sum_{m=0}^{n} \sum_{\mu=1}^{\infty} \left( A \cdot t_m^k \right) \cdot e^{-B \cdot t_n},$$

(24)

which is equivalent to (12).
Figure 1: The parameters effect on the adoption rate, as well as the effect on the broadband penetration.

Figure 2: The influence of parameters on the model performance.
Consequently, parameter $D$ relates to the time period of the growth as well as to the slope of the adoption rate curve (Figure 1). Figure 2 shows the influence of the model parameters on its performance.

4. Market Dataset

4.1. Market Overview. National Telecommunications and Post Commission (EETT) is the regulatory authority in Greece. EETT’s objective is the development of the broadband market in Greece, while at the same time enforcing the fair competition and consumer protection [18]. Digital Subscriber Line (DSL) technology is the predominant of the fixed broadband market. The DSL coverage is available in terms of percentage of population. According to the European Commission (EU), the DSL availability covers 100% of urban areas, while in rural areas is 50% [19]. According to the OECD Broadband Portal December 2008 data [20], 95% of fixed broadband connections in Greece are based on xDSL technologies. Broadband cable access is not available in Greece. Mobile broadband is at an early growth stage, lower than the EU average [19–24]. It should be noted that EETT’s data refer only to fixed broadband connections, without taking into account the penetration of broadband mobile networks.

4.2. Dataset. The actual data used in this short analysis concerns the quarterly total broadband connections and the percentage of broadband connections per 100 inhabitants in Greece, from December 2004 until June 2010 (see Figure 3) [25]. The construction of Figure 3 is based on data of the National Telecommunications and Post Commission (EETT) [21–23], as well as data of the Observatory for the Greek Information Society [24].

5. Fitting and Forecasting Method

Regression analysis is used to fit sampled data to a model. Curve fitting is done using the ordinary least squares method (OLS) [12]. The objective of ordinary least squares method is to minimize the sum of squared error (SSE) between data points $(t_i, O(t_i))$ and model evaluated points $(t_i, Y(t_i; a_1, a_2, \ldots, a_k))$:

$$SSE_{OLS} = \sum_{i=1}^{n} [O(t_i) - Y(t_i; a_1, a_2, \ldots, a_k)]^2,$$

where $Y(t_i; a_1, a_2, \ldots, a_k)$ is a time function with a set of $k$ parameters $\{a_i\}$ and $n > k$ known data points $(t_i, O(t_i))$ [12].

In forecasting, parameter estimation is usually focused on the time interval near the last observed data points. Thus, the weighted least squares method is used [12]. The weighted least squares method is to minimize the weighted sum of squared error (SSE) between points of the dataset $(t_i, O(t_i))$ and model estimated points $(t_i, Y(t_i; a_1, a_2, \ldots, a_k))$:

$$SSE_{WLS} = \sum_{i=1}^{n} w_i \cdot [O(t_i) - Y(t_i; a_1, a_2, \ldots, a_k)]^2,$$

where $w_i$ is the weight $1/\sigma_i$ and $\sigma_i^2$ is the variance of $i$th observation [12].
Table 1: Results of the fitting process for the models.

| Model                        | Mean Absolute Error (MAE) | R-Square | Correlation | Mean Absolute Percentage Error (MAPE) | Mean Squared Error (MSE) | Root Mean Squared Error (RMSE) | Sum of Squared Error (SSE) |OLS |
|------------------------------|---------------------------|----------|-------------|--------------------------------------|--------------------------|--------------------------------|------------------------------|
| Forecasting model            | 0.00141239                | 0.99897498 | 0.999285392 | 0.035314                              | 3.885E-06                | 1.97E-03                        | 8.936E-05                   |
| Gompertz with constant       | 0.00155774                | 0.99885993 | 0.99921341  | 0.039917                              | 4.321E-06                | 2.08E-03                        | 9.939E-05                   |
| Gompertz                     | 0.00168856                | 0.99878231 | 0.999155808 | 0.066631                              | 4.665E-06                | 2.16E-03                        | 1.073E-04                   |
| Richards                     | 0.00169607                | 0.99856237 | 0.999070053 | 0.084240                              | 5.523E-06                | 2.35E-03                        | 1.270E-04                   |
| Box Cox                      | 0.00192879                | 0.99851900 | 0.999058878 | 0.088561                              | 5.705E-06                | 2.39E-03                        | 1.312E-04                   |
| FLOG                         | 0.00220999                | 0.99810868 | 0.99880778  | 0.091428                              | 7.168E-06                | 2.68E-03                        | 1.649E-04                   |
| Logistic with constant       | 0.00220926                | 0.99810868 | 0.9988074   | 0.091236                              | 7.169E-06                | 2.68E-03                        | 1.649E-04                   |
| Bass                         | 0.00296358                | 0.99696321 | 0.998073643 | 0.113653                              | 1.227E-05                | 3.50E-03                        | 2.823E-04                   |

Figure 4: Residuals per model against time.

6. Results

6.1. Fitting Results. According to the Sum of Squared Error (SSE)_{OLS}, the fitting curves are sorted in Table 1.

The forecasting model shows the best fitting performance. It should be noted that the statistical indices concern the whole dataset. The fitting performance of the forecasting model improves its performance for any subset of the dataset, especially for the latest data. Analytically, the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the Mean Absolute Percentage Error (MAPE) are statistics for evaluating the overall quality of a regression model [26]. The Mean Absolute Error (MAE) is the average of the absolute value of the residuals. We observe that the MAE of the empirical model is the smallest concerning the whole set of models. It should be taken into account that this reduction is observed mainly in the last data. MAE is very similar to RMSE, but less sensitive to large errors. It can also be seen that the Root Mean Squared Error (RMSE), which is the square root of the average squared distance of a data point from the fitted line, is the smallest. The Mean Absolute Percentage Error (MAPE) is the percentage...
Fitting performance

Table 2: Results of the forecasting process for the models (forecasting period: one year).

<table>
<thead>
<tr>
<th>Model</th>
<th>Forecasting performance (forecasting period: one year)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Absolute Error (MAE)</td>
<td>Mean Absolute Percentage Error (MAPE)</td>
<td>Mean Squared Error (MSE)</td>
<td>Root Mean Squared Error (RMSE)</td>
</tr>
<tr>
<td>Forecasting model</td>
<td>0.001812971</td>
<td>0.010961673</td>
<td>4.2351E – 06</td>
<td>0.002057930</td>
</tr>
<tr>
<td>Gompertz</td>
<td>0.002374115</td>
<td>0.014203962</td>
<td>7.3493E – 06</td>
<td>0.002710962</td>
</tr>
<tr>
<td>Richards</td>
<td>0.002589747</td>
<td>0.015089135</td>
<td>8.2087E – 06</td>
<td>0.002865074</td>
</tr>
<tr>
<td>Gompertz with constant</td>
<td>0.002907623</td>
<td>0.016882766</td>
<td>9.8438E – 06</td>
<td>0.003137484</td>
</tr>
<tr>
<td>Bass</td>
<td>0.003035126</td>
<td>0.017946140</td>
<td>1.1443E – 05</td>
<td>0.003382720</td>
</tr>
<tr>
<td>Logistic with constant</td>
<td>0.003061642</td>
<td>0.018094383</td>
<td>1.1615E – 05</td>
<td>0.003408084</td>
</tr>
<tr>
<td>Box Cox</td>
<td>0.003216461</td>
<td>0.018936105</td>
<td>1.2360E – 05</td>
<td>0.003515620</td>
</tr>
<tr>
<td>FLOG</td>
<td>0.003306149</td>
<td>0.019436138</td>
<td>1.2983E – 05</td>
<td>0.003603245</td>
</tr>
<tr>
<td>Logistic</td>
<td>0.009994163</td>
<td>0.056669925</td>
<td>1.3492E – 04</td>
<td>0.011615528</td>
</tr>
</tbody>
</table>

of the summarized absolute error divided by the summarized actual values of the observation. The MAPE indicator of the empirical model is also the smallest. Figure 4 presents the errors of the models in time, namely, residuals per model.

It can be seen that the MAE of the forecasting model is the smallest. In addition, Gompertz (with constant), Gompertz, and Richards’ models show a good behaviour. The aforementioned performance is improved on the time interval near the last observed data points. In Figure 5, the fitting model plots in time, December 2004 until June 2010, is shown.

The overall fitting performance of the models is satisfactory, as is the correlation coefficient between fitted and actual data of the models. This coefficient varies between 0.999285392 (forecasting model) and 0.998073643 (logistic model). In case the parameters of the models are calculated by the OLS regression method, the R squared ($R^2$) indicator which is the square of the correlation coefficient depends on the number of the parameters. So, the indicator’s range varies.
6.2. Forecasting Results. The parameters are estimated by the means of regression analysis (WLS method). The dataset consists of 23 data points (December 2004 until June 2010). The forecasting period is one year, June 2009 until June 2010 (after 19th data point). The statistical indices concerning the forecasting performance of the models are presented in Table 2.

In the literature, MAPE is a reliable indicator for evaluating the prediction performance of the models. The proposed model performs better than the others. Specifically, the empirical model achieves well enough indicators (MAPE, MAE, and MSE). In general, it could be argued that the models that achieve good indicators, especially MAPE, are the Richards and Gompertz family.

In Figure 6, the aforementioned performance is graphically presented, for a forecasting period of one year.

6.3. Future Trends. The performance of the models using the dataset (23 data points) is presented here. The WLS method was also applied here. The parameters of the forecasting model are presented in Table 3.

According to the MAPE indicator, the forecasting curves are sorted in Table 4. According to the MAPE indicator, the forecasting model, Gompertz with constant, Richards and Gompertz perform better than the others. According to SSEWLS, the models with the best performance are the forecasting model, FLOG, Logistics (with constant), and Bass. The estimation of our model for the broadband penetration in Greece on June 2012 (two years ahead) is 22.09% approximately. The most optimistic estimation concerns the Gompertz model (BB penetration 22.34%), and the most pessimistic estimation is done by Logistic with 19.46%. Also for forecasting period of one year, the empirical model estimates the BB penetration 20.73%, Gompertz 20.85%, Gompertz (constant) 20.75%, and Richards 20.34% (Figure 7).

7. Future Implementation of the Empirical Model

It is mentioned that the adoption of an innovative technology by society follows the sigmoid curve. So, the
Table 4: Results of the forecasting process for the models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean Absolute Error (MAE)</th>
<th>R-Square</th>
<th>Correlation</th>
<th>Mean Absolute Percentage Error (MAPE)</th>
<th>Mean Squared Error (MSE)</th>
<th>Root Mean Squared Error (RMSE)</th>
<th>Sum of Squared Error (SSE)_{WLS}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecasting Model</td>
<td>0.00150121</td>
<td>0.99877253</td>
<td>0.99920491</td>
<td>0.032060</td>
<td>4.719E − 06</td>
<td>2.17E − 03</td>
<td>0.049023031</td>
</tr>
<tr>
<td>Gompertz with constant</td>
<td>0.00158061</td>
<td>0.99876359</td>
<td>0.99920045</td>
<td>0.033855</td>
<td>4.346E − 06</td>
<td>2.08E − 03</td>
<td>0.058116463</td>
</tr>
<tr>
<td>Richards</td>
<td>0.00186277</td>
<td>0.99834027</td>
<td>0.998929485</td>
<td>0.040750</td>
<td>5.837E − 06</td>
<td>2.42E − 03</td>
<td>0.068116161</td>
</tr>
<tr>
<td>Gompertz</td>
<td>0.00160523</td>
<td>0.99874607</td>
<td>0.999194834</td>
<td>0.047587</td>
<td>4.759E − 06</td>
<td>2.18E − 03</td>
<td>0.099948481</td>
</tr>
<tr>
<td>Box Cox</td>
<td>0.00200287</td>
<td>0.99844119</td>
<td>0.998967742</td>
<td>0.062505</td>
<td>6.161E − 06</td>
<td>2.48E − 03</td>
<td>0.054860705</td>
</tr>
<tr>
<td>FLOG</td>
<td>0.00231788</td>
<td>0.99800684</td>
<td>0.99867706</td>
<td>0.084474</td>
<td>8.561E − 06</td>
<td>2.93E − 03</td>
<td>0.056633745</td>
</tr>
<tr>
<td>Bass</td>
<td>0.00231846</td>
<td>0.99800583</td>
<td>0.998677</td>
<td>0.084541</td>
<td>8.569E − 06</td>
<td>2.93E − 03</td>
<td>0.056633711</td>
</tr>
<tr>
<td>Logistic with constant</td>
<td>0.00323439</td>
<td>0.99696693</td>
<td>0.998468875</td>
<td>0.111882</td>
<td>1.385E − 05</td>
<td>3.72E − 03</td>
<td>0.274783412</td>
</tr>
<tr>
<td>Logistic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7: The overall forecasting performance of the models.

The empirical model could be implemented for a general-purpose time-series forecasting that follows the sigmoid curve. The implementation to different markets would give, using regression analysis, different parameters estimation.

The time horizon of the forecast depends on the dataset density. Generally, it should be noted that the reliability of the diffusion models depends on the number of the time series data. This principle governs the empirical model. So the implementation of the model should be chosen based on this principle.

Specifically, the coverage of the FTTH (Fiber To The Home) technology is rapidly growing, according to OECD [27]. So, a suggestion of future implementation of the model (when the historical data are satisfactory in number) would be the forecast of the FTTH technology penetration in a
country or in a geographical sector (e.g., Europe, Asia, America, etc.)

8. Conclusion

In this paper, a new short-term forecasting model was introduced, concerning the overall broadband penetration. The forecasting model exhibited a well-fitting performance to the observed data. The residuals, as well as the RMSE, MAE, and MAPE indicators of the empirical model were satisfactory. The statistical indicators, concerning the forecasting behaviour of the model, namely, MAE, RMSE, and MAPE, showed satisfactory results, also.

Future study about the performance of the empirical model for the broadband penetration, in different markets, should be considered. The comparison of the behaviour of the proposed model in different markets can offer a better estimation of the model’s parameters and the correlation with financial and social indices.

Acknowledgment

The authors wish to express their acknowledgments to Professor Luis Carlos Rabelo, University of Central Florida, USA, for his constructive comments and suggestions, which helped to improve the quality of this paper.

References

[27] OECD Broadband Portal, Press release, http://www.oecd.org/document/54/0,3746,en_264934225_395756701_1_1_1_1,00.html.